



# SUBSTITUTE SEQUENCE LISTING

<110> David Baltimore et al.

<120> NUCLEAR FACTORS ASSOCIATED WITH TRANSCRIPTIONAL REGULATION

<130> APBI-P05-035

<140> 10/037341

<141> 2002-01-04

<150> 08/464364

<151> 1995-06-05

<160> 60

<170> PatentIn version 3.1

<210> 1

<211> 12

<212> DNA

<213> Homo sapiens

<400> 1

tggggattcc ca

12

<210> 2

<211> 11

<212> DNA

<213> Mus musculus

<400> 2

ggggactttc c

11

<210> 3

<211> 11

<212> DNA

<213> HIV

<400> 3

agggactttc c

11

<210> 4

<211> 10

<212> DNA

<213> Homo sapiens

<400> 4

ggggatttcc

10

<210> 5

<211> 10

<212> DNA

<213> Homo sapiens

<400> 5	10
gggaaattcc	
<210> 6	
<211> 10	
<212> DNA	
<213> CMV	
<400> 6	10
gggactttcc	
<210> 7	
<211> 10	
<212> DNA	
<213> Homo sapiens	
<400> 7	10
gggacttccc	
<210> 8	
<211> 10	
<212> DNA	
<213> Homo sapiens	
<400> 8	10
gggatttcac	
<210> 9	
<211> 10	
<212> DNA	
<213> Mus musculus	
<400> 9	10
ggggattcct	
<210> 10	
<211> 10	
<212> DNA	
<213> Homo sapiens	
<400> 10	10
gggaatctcc	
<210> 11	
<211> 10	
<212> DNA	
<213> Homo sapiens	
<400> 11	10
gggattcccc	
<210> 12	
<211> 10	
<212> DNA	
<213> Artificial Sequence	

<220>  
<223> Consensus sequence

<220>  
<221> misc\_feature  
<222> (5)..(5)  
<223> n=a, c or t

<220>  
<221> misc\_feature  
<222> (9)..(9)  
<223> n=a, c or t

<400> 12  
gggrntyync

10

<210> 13  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic binding site

<400> 13  
aaatccccta aaacgaggga taaa

24

<210> 14  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Oligonucleotide for pUCMHCI construction

<400> 14  
gatccggctg gggattcccc atct

24

<210> 15  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Oligonucleotide for pUCMHCI construction

<400> 15  
gatccggctg cggattccca atct

24

<210> 16  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
 <223> Oligonucleotide for pUCmhci construction  
  
 <400> 16 24  
 gccgaccct aaggggtaga ctag  
  
 <210> 17  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Oligonucleotide for pUCmhcI construction  
  
 <400> 17 24  
 gccgacgcct aaggggtaga ctag  
  
 <210> 18  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Oligonucleotide for Plasmid pBS-ATG construction  
  
 <400> 18 27  
 tgcacacccat ggccatcgat atcgatc  
  
 <210> 19  
 <211> 10  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Mutant showing slight competition for binding of lysogen protein  
  
 <400> 19 10  
 tcatttccat  
  
 <210> 20  
 <211> 34  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Nucleotide sequence of the PRDIIx2  
  
 <400> 20 34  
 gatctgtggg aaattccgtg ggaaattccg gatc  
  
 <210> 21  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> KB wild type oligonucleotide

<400> 21 30  
 tcgacagagg ggactttccg agaggctcga

<210> 22  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> KB mutant oligonucleotide

<400> 22 30  
 tcgacagaat tcactttcca ggaggctcga

<210> 23  
 <211> 31  
 <212> DNA  
 <213> Homo sapiens

<400> 23 31  
 gcttcttaat aatttgcata ccctcactgc a

<210> 24  
 <211> 24  
 <212> DNA  
 <213> Homo sapiens

<400> 24 24  
 tcttaataat ttgcataccc tcac

<210> 25  
 <211> 24  
 <212> DNA  
 <213> Homo sapiens

<400> 25 24  
 cgcacatgat ttgcatactc atga

<210> 26  
 <211> 24  
 <212> DNA  
 <213> Homo sapiens

<400> 26 24  
 cctgggtaat ttgcatttct aaaa

<210> 27  
 <211> 75  
 <212> DNA  
 <213> Homo sapiens

<400> 27 60  
 ctaactgctt cttaataatt tgcataccct cactgcatcg ccttggggac ttctttatat  
 aacagtcaaa catat 75

<210> 28  
 <211> 22  
 <212> DNA  
 <213> Homo sapiens

<400> 28  
 aattaccag gtggtgtttt gc 22

<210> 29  
 <211> 22  
 <212> DNA  
 <213> Homo sapiens

<400> 29  
 agcagctcat gtggcaaggc ta 22

<210> 30  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> EBNA-1 Consensus binding site

<400> 30  
 agattaggat agcatatgct acccagatat 30

<210> 31  
 <211> 13  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> DNA probe

<400> 31  
 tggggattcc cca 13

<210> 32  
 <211> 13  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> DNA probe

<400> 32  
 tgcggattcc caa 13

<210> 33  
 <211> 14  
 <212> DNA  
 <213> Artificial Sequence

<220>		
<223>	DNA probe	
<400>	33	14
	aggggacttt ccgg	
<210>	34	
<211>	12	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	DNA probe	
<400>	34	12
	aaatactttc cg	
<210>	35	
<211>	13	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	DNA probe	
<400>	35	13
	tggggacttt cca	
<210>	36	
<211>	13	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	DNA probe	
<400>	36	13
	tggggacttt cca	
<210>	37	
<211>	13	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	DNA probe	
<400>	37	13
	aagggacttt ccg	
<210>	38	
<211>	1717	

<212> DNA  
<213> Homo sapiens

<220>  
<221> CDS  
<222> (67)..(1464)  
<223>

<400> 38  
ctggggcccc cagagagggt ggggagatga cacagttggt cccccagccc tggcggggcg 60  
ggcagc atg gtt cac tcc agc atg ggg gct cca gaa ata aga atg tct 108  
Met Val His Ser Ser Met Gly Ala Pro Glu Ile Arg Met Ser  
1 5 10  
aag ccc ctg gag gcc gag aag caa ggt ctg gac tcc cca tca gag cac 156  
Lys Pro Leu Glu Ala Glu Lys Gln Gly Leu Asp Ser Pro Ser Glu His  
15 20 25 30  
aca gac acc gaa aga aat gga cca gac act aat cat cag aac ccc caa 204  
Thr Asp Thr Glu Arg Asn Gly Pro Asp Thr Asn His Gln Asn Pro Gln  
35 40 45  
aat aag acc tcc cca ttc tcc gtg tcc cca act ggc ccc agt aca aag 252  
Asn Lys Thr Ser Pro Phe Ser Val Ser Pro Thr Gly Pro Ser Thr Lys  
50 55 60  
atc aag gct gaa gac ccc agt ggc gat tca gcc cca gca gca ccc ctg 300  
Ile Lys Ala Glu Asp Pro Ser Gly Asp Ser Ala Pro Ala Ala Pro Leu  
65 70 75  
ccc cct cag ccg gcc cag cct cat ctg ccc cag gcc caa ctc atg ttg 348  
Pro Pro Gln Pro Ala Gln Pro His Leu Pro Gln Ala Gln Leu Met Leu  
80 85 90  
acg ggc agc cag cta gct ggg gac ata cag cag ctc ctc cag ctc cag 396  
Thr Gly Ser Gln Leu Ala Gly Asp Ile Gln Gln Leu Leu Gln Leu Gln  
95 100 105 110  
cag ctg gtg ctt gtg cca ggc cac cac ctc cag cca cct gct cag ttc 444  
Gln Leu Val Leu Val Pro Gly His His Leu Gln Pro Pro Ala Gln Phe  
115 120 125  
ctg cta ccg cag gcc cag cag agc cag cca ggc ctg cta ccg aca cca 492  
Leu Leu Pro Gln Ala Gln Gln Ser Gln Pro Gly Leu Leu Pro Thr Pro  
130 135 140  
aat cta ttc cag cta cct cag caa acc cag gga gct ctt ctg acc tcc 540  
Asn Leu Phe Gln Leu Pro Gln Gln Thr Gln Gly Ala Leu Leu Thr Ser  
145 150 155  
cag ccc cgg gcc ggg ctt ccc aca cag gcc gtg acc cgc cct acg ctg 588  
Gln Pro Arg Ala Gly Leu Pro Thr Gln Ala Val Thr Arg Pro Thr Leu  
160 165 170  
ccc gac ccg cac ctc tcg cac ccg cag ccc ccc aaa tgc ttg gag cca 636  
Pro Asp Pro His Leu Ser His Pro Gln Pro Pro Lys Cys Leu Glu Pro  
175 180 185 190  
cca tcc cac ccc gag gag ccc agt gat ctg gag gag ctg gag caa ttc 684  
Pro Ser His Pro Glu Glu Pro Ser Asp Leu Glu Glu Leu Glu Gln Phe  
195 200 205  
gcc cgc acc ttc aag caa cgc cgc atc aag ctg ggc ttc acg cag ggt 732  
Ala Arg Thr Phe Lys Gln Arg Arg Ile Lys Leu Gly Phe Thr Gln Gly  
210 215 220  
gat gtg ggc ctg gcc atg ggc aag ctc tac ggc aac gac ttc agc cag 780  
Asp Val Gly Leu Ala Met Gly Lys Leu Tyr Gly Asn Asp Phe Ser Gln  
225 230 235  
acg acc att tcc cgc ttc gag gcc ctc aac ctg agc ttc aag aac atg 828  
Thr Thr Ile Ser Arg Phe Glu Ala Leu Asn Leu Ser Phe Lys Asn Met



240	245	250	
tgc aaa ctc aag ccc ctc ctg gag aag tgg ctc aac gat gca gag act			876
Cys Lys Leu Lys Pro Leu Leu Glu Lys Trp Leu Asn Asp Ala Glu Thr			
255	260	265	270
atg tct gtg gac tca agc ctg ccc agc ccc aac cag ctg agc agc ccc			924
Met Ser Val Asp Ser Ser Leu Pro Ser Pro Asn Gln Leu Ser Ser Pro			
275	280	285	
agc ctg ggt ttc gac ggc ctg ccc ggc cgg aga cgc aag aag agg acc			972
Ser Leu Gly Phe Asp Gly Leu Pro Gly Arg Arg Arg Lys Lys Arg Thr			
290	295	300	
agc atc gag aca aac gtc cgc ttc gcc tta gag aag agt ttt cta gcg			1020
Ser Ile Glu Thr Asn Val Arg Phe Ala Leu Glu Lys Ser Phe Leu Ala			
305	310	315	
aac cag aag cct acc tca gag gag atc ctg ctg atc gcc gag cag ctg			1068
Asn Gln Lys Pro Thr Ser Glu Glu Ile Leu Leu Ile Ala Glu Gln Leu			
320	325	330	
cac atg gag aag gaa gtg atc cgc gtc tgg ttc tgc aac cgg cgc cag			1116
His Met Glu Lys Glu Val Ile Arg Val Trp Phe Cys Asn Arg Arg Gln			
335	340	345	350
aag gag aaa cgc atc aac ccc tgc agt gcg gcc ccc atg ctg ccc agc			1164
Lys Glu Lys Arg Ile Asn Pro Cys Ser Ala Ala Pro Met Leu Pro Ser			
355	360	365	
cca ggg aag ccg gcc agc tac agc ccc cat atg gtc aca ccc caa ggg			1212
Pro Gly Lys Pro Ala Ser Tyr Ser Pro His Met Val Thr Pro Gln Gly			
370	375	380	
ggc gcg ggg acc tta ccg ttg tcc caa gct tcc agc agt ctg agc aca			1260
Gly Ala Gly Thr Leu Pro Leu Ser Gln Ala Ser Ser Ser Leu Ser Thr			
385	390	395	
aca gtt act acc tta tcc tca gct gtg ggg acg ctc cac ccc agc cgg			1308
Thr Val Thr Thr Leu Ser Ser Ala Val Gly Thr Leu His Pro Ser Arg			
400	405	410	
aca gct gga ggg ggt ggg ggc ggg ggc ggg gct gcg ccc ccc ctc aat			1356
Thr Ala Gly Gly Gly Gly Gly Gly Gly Gly Ala Ala Pro Pro Leu Asn			
415	420	425	430
tcc atc ccc tct gtc act ccc cca ccc ccg gcc acc acc aac agc aca			1404
Ser Ile Pro Ser Val Thr Pro Pro Pro Pro Ala Thr Thr Asn Ser Thr			
435	440	445	
aac ccc agc cct caa ggc agc cac tcg gct atc ggc ttg tca ggc ctg			1452
Asn Pro Ser Pro Gln Gly Ser His Ser Ala Ile Gly Leu Ser Gly Leu			
450	455	460	
aac ccc agc acg gggtaagtgg gtgcacgtgg gaagctgtgg ggagaagcag			1504
Asn Pro Ser Thr			
465			
ggtcgctgct gcttctaggg tggggagcgg caccacagtt atgttggcag gtccctgccc			1564
ctgctaatagc ctctgctttg cctcttgacg aagcacaatg gtgggggttg gctccggctg			1624
agtccagccc tcatgagcaa caaccctttg gccactatcc aaggtgcgtg ctgcctcatg			1684
tcacacccat cgtcaccagc cccggaattc gag			1717

<210> 39  
 <211> 466  
 <212> PRT  
 <213> Homo sapiens

<400> 39  
 Met Val His Ser Ser Met Gly Ala Pro Glu Ile Arg Met Ser Lys Pro  
 1 5 10 15

Leu Glu Ala Glu Lys Gln Gly Leu Asp Ser Pro Ser Glu His Thr Asp  
 20 25 30  
 Thr Glu Arg Asn Gly Pro Asp Thr Asn His Gln Asn Pro Gln Asn Lys  
 35 40 45  
 Thr Ser Pro Phe Ser Val Ser Pro Thr Gly Pro Ser Thr Lys Ile Lys  
 50 55 60  
 Ala Glu Asp Pro Ser Gly Asp Ser Ala Pro Ala Ala Pro Leu Pro Pro  
 65 70 75 80  
 Gln Pro Ala Gln Pro His Leu Pro Gln Ala Gln Leu Met Leu Thr Gly  
 85 90 95  
 Ser Gln Leu Ala Gly Asp Ile Gln Gln Leu Leu Gln Leu Gln Gln Leu  
 100 105 110  
 Val Leu Val Pro Gly His His Leu Gln Pro Pro Ala Gln Phe Leu Leu  
 115 120 125  
 Pro Gln Ala Gln Gln Ser Gln Pro Gly Leu Leu Pro Thr Pro Asn Leu  
 130 135 140  
 Phe Gln Leu Pro Gln Gln Thr Gln Gly Ala Leu Leu Thr Ser Gln Pro  
 145 150 155 160  
 Arg Ala Gly Leu Pro Thr Gln Ala Val Thr Arg Pro Thr Leu Pro Asp  
 165 170 175  
 Pro His Leu Ser His Pro Gln Pro Pro Lys Cys Leu Glu Pro Pro Ser  
 180 185 190  
 His Pro Glu Glu Pro Ser Asp Leu Glu Glu Leu Glu Gln Phe Ala Arg  
 195 200 205  
 Thr Phe Lys Gln Arg Arg Ile Lys Leu Gly Phe Thr Gln Gly Asp Val  
 210 215 220  
 Gly Leu Ala Met Gly Lys Leu Tyr Gly Asn Asp Phe Ser Gln Thr Thr  
 225 230 235 240  
 Ile Ser Arg Phe Glu Ala Leu Asn Leu Ser Phe Lys Asn Met Cys Lys  
 245 250 255  
 Leu Lys Pro Leu Leu Glu Lys Trp Leu Asn Asp Ala Glu Thr Met Ser  
 260 265 270  
 Val Asp Ser Ser Leu Pro Ser Pro Asn Gln Leu Ser Ser Pro Ser Leu  
 275 280 285  
 Gly Phe Asp Gly Leu Pro Gly Arg Arg Arg Lys Lys Arg Thr Ser Ile  
 290 295 300  
 Glu Thr Asn Val Arg Phe Ala Leu Glu Lys Ser Phe Leu Ala Asn Gln  
 305 310 315 320

Lys Pro Thr Ser Glu Glu Ile Leu Leu Ile Ala Glu Gln Leu His Met  
325 330 335

Glu Lys Glu Val Ile Arg Val Trp Phe Cys Asn Arg Arg Gln Lys Glu  
340 345 350

Lys Arg Ile Asn Pro Cys Ser Ala Ala Pro Met Leu Pro Ser Pro Gly  
355 360 365

Lys Pro Ala Ser Tyr Ser Pro His Met Val Thr Pro Gln Gly Gly Ala  
370 375 380

Gly Thr Leu Pro Leu Ser Gln Ala Ser Ser Ser Leu Ser Thr Thr Val  
385 390 395 400

Thr Thr Leu Ser Ser Ala Val Gly Thr Leu His Pro Ser Arg Thr Ala  
405 410 415

Gly Gly Gly Gly Gly Gly Gly Gly Ala Ala Pro Pro Leu Asn Ser Ile  
420 425 430

Pro Ser Val Thr Pro Pro Pro Pro Ala Thr Thr Asn Ser Thr Asn Pro  
435 440 445

Ser Pro Gln Gly Ser His Ser Ala Ile Gly Leu Ser Gly Leu Asn Pro  
450 455 460

Ser Thr  
465

<210> 40  
<211> 277  
<212> PRT  
<213> Homo sapiens

<400> 40  
Cys Gly Pro Gly His Gly Gln Ala Leu Arg Gln Arg Leu Gln Pro Asp  
1 5 10 15

Asp His Phe Pro Leu Arg Gly Pro Gln Pro Glu Leu Gln Glu His Val  
20 25 30

Gln Thr Gln Ala Pro Pro Gly Glu Val Ala Gln Arg Cys Arg Asp Tyr  
35 40 45

Val Cys Gly Leu Lys Pro Ala Gln Pro Gln Pro Ala Glu Gln Pro Gln  
50 55 60

Pro Gly Phe Arg Ala Cys Met Pro Glu Thr Gln Glu Glu Asp Gln Met  
65 70 75 80

Arg Asp Lys Lys Pro Leu Arg Leu Arg Glu Glu Phe Ser Ser Glu Pro  
85 90 95

Glu Ala Tyr Leu Arg Gly Asp Pro Ala Asp Arg Arg Ala Ala Ala His  
100 105 110

Gly Glu Gly Ser Asp Pro Arg Leu Val Leu Gln Pro Ala Pro Glu Gly  
 115 120 125  
 Glu Thr His Gln Pro Leu Gln Cys Gly Pro His Ala Ala Gln Pro Arg  
 130 135 140  
 Glu Ala Gly Gln Leu Gln Pro Pro Tyr Gly His Thr Pro Ala Gly Arg  
 145 150 155 160  
 Gly Asp Leu Thr Val Val Pro Ser Phe Gln Gln Ser Glu His Asn Ser  
 165 170 175  
 Tyr Tyr Leu Ile Leu Ser Cys Gly Asp Ala Pro Pro Gln Pro Asp Ser  
 180 185 190  
 Asn Met Gly Trp Gly Met Gly Arg Gly Cys Ala Pro Pro Gln Phe His  
 195 200 205  
 Pro Leu Cys His Ser Pro Thr Pro Gly His Asn Gln Gln His Lys Pro  
 210 215 220  
 Gln Pro Ser Arg Gln Pro Leu Gly Tyr Met Leu Val Ala Pro Glu Pro  
 225 230 235 240  
 Gln Asn Gly Val Ser Gly Cys Thr Trp Glu Ala Val Gly Arg Ser Arg  
 245 250 255  
 Val Ala Ala Ala Ser Arg Val Gly Ser Gly Thr Pro Val Met Leu Ala  
 260 265 270  
 Gly Pro Cys Pro Cys  
 275

<210> 41  
 <211> 437  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (1)..(90)  
 <223>

<400> 41  
 cct caa ggc agc cac tcg gct atc ggc ttg tca ggc ctg aac ccc agc 48  
 Pro Gln Gly Ser His Ser Ala Ile Gly Leu Ser Gly Leu Asn Pro Ser  
 1 5 10 15  
 acg ggc cct ggc ctc tgg tgg aac cct gcc cct tac cag cct 90  
 Thr Gly Pro Gly Leu Trp Trp Asn Pro Ala Pro Tyr Gln Pro  
 20 25 30  
 tgatggcagc gggaatctgg tgctgggggc agccggtgca gccccgggga gccctggcct 150  
 ggtgacctcg ccgctcttct tgaatcatgc tgggctgccc ctgctcagca ccccgcttgg 210  
 tgtgggcctg gtctcagcag cggctgcggg tgtggcagcc tccatctcca gcaagtctcc 270  
 tggcctctcc tcctcatcct cttcatcctc atcctcctcc tcctccactt gcagcgagac 330  
 ggcagcacag accctggagg tccagggggg cccgaggcag ggtccaaacc tgagtgaggg 390  
 ccagccatgc ctcccctccc attcctctgg tcctgcccc ggaattc 437

<210> 42  
<211> 30  
<212> PRT  
<213> Homo sapiens

<400> 42  
Pro Gln Gly Ser His Ser Ala Ile Gly Leu Ser Gly Leu Asn Pro Ser  
1 5 10 15  
Thr Gly Pro Gly Leu Trp Trp Asn Pro Ala Pro Tyr Gln Pro  
20 25 30

<210> 43  
<211> 50  
<212> PRT  
<213> Homo sapiens

<400> 43  
Ser Ala Gln Pro Leu Gly Tyr Arg Leu Val Met Pro Glu Pro Gln Met  
1 5 10 15  
Gly Pro Asn Pro Leu Val Glu Pro Cys Pro Leu Pro Ala Leu Met Ala  
20 25 30  
Ala Gly Ile Trp Cys Trp Gly Gln Pro Val Gln Pro Arg Gly Ala Leu  
35 40 45

Ala Trp  
50

<210> 44  
<211> 62  
<212> PRT  
<213> Homo sapiens

<400> 44  
Arg Arg Lys Lys Arg Thr Ser Ile Glu Thr Asn Val Arg Phe Ala Leu  
1 5 10 15  
Glu Lys Ser Phe Leu Ala Asn Gln Lys Pro Thr Ser Glu Glu Ile Leu  
20 25 30  
Leu Ile Ala Glu Gln Leu His Met Glu Lys Glu Val Ile Arg Val Trp  
35 40 45  
Phe Cys Asn Arg Arg Gln Lys Glu Lys Arg Ile Asn Pro Cys  
50 55 60

<210> 45  
<211> 57  
<212> PRT  
<213> Homo sapiens

<400> 45  
Ser Pro Lys Gly Lys Ser Ser Ile Ser Pro Gln Ala Arg Ala Phe Leu  
1 5 10 15

Glu Gln Val Phe Arg Arg Lys Gln Ser Leu Asn Ser Lys Glu Lys Glu  
20 25 30

Glu Val Ala Lys Lys Cys Gly Ile Thr Pro Leu Gln Val Arg Val Trp  
35 40 45

Phe Ile Asn Lys Arg Met Arg Ser Lys  
50 55

<210> 46

<211> 59

<212> PRT

<213> Homo sapiens

<400> 46

Lys Pro Tyr Arg Gly His Arg Phe Thr Lys Glu Asn Val Arg Ile Leu  
1 5 10 15

Glu Ser Trp Phe Ala Lys Asn Pro Tyr Leu Asp Thr Lys Gly Leu Glu  
20 25 30

Asn Leu Met Asn Thr Ser Leu Ser Arg Ile Gln Ile Lys Asn Trp Val  
35 40 45

Ser Asn Arg Arg Arg Lys Glu Lys Thr Ile Thr  
50 55

<210> 47

<211> 60

<212> PRT

<213> Homo sapiens

<400> 47

Gln Arg Pro Lys Arg Thr Arg Ala Lys Gly Glu Ala Leu Asp Val Leu  
1 5 10 15

Lys Arg Lys Phe Glu Ile Asn Pro Thr Pro Ser Leu Val Glu Arg Lys  
20 25 30

Lys Ile Ser Asp Leu Ile Gly Met Pro Glu Lys Asn Val Arg Ile Trp  
35 40 45

Phe Gln Asn Arg Arg Ser Lys Glu Arg Arg Leu Lys  
50 55 60

<210> 48

<211> 60

<212> PRT

<213> Homo sapiens

<400> 48

Arg Arg Gly Pro Arg Thr Thr Ile Lys Gln Asn Gln Leu Asp Val Leu  
1 5 10 15

Asn Glu Met Phe Ser Asn Thr Pro Lys Pro Ser Lys His Ala Arg Ala  
20 25 30

Lys Leu Ala Leu Glu Thr Gly Leu Ser Met Arg Val Ile Gln Val Trp  
35 40 45

Phe Gln Asn Arg Arg Ser Lys Glu Arg Arg Leu Lys  
50 55 60

<210> 49  
<211> 60  
<212> PRT  
<213> Homo sapiens

<400> 49  
Ser Lys Lys Gln Arg Val Leu Phe Ser Glu Glu Gln Lys Glu Ala Leu /  
1 5 10 15

Arg Leu Ala Phe Ala Leu Asp Pro Tyr Pro Asn Val Gly Thr Ile Glu  
20 25 30

Phe Leu Ala Asn Glu Leu Gly Leu Ala Thr Arg Thr Ile Thr Asn Trp  
35 40 45

Phe His Asn His Arg Met Arg Leu Lys Gln Gln Val  
50 55 60

<210> 50  
<211> 60  
<212> PRT  
<213> Homo sapiens

<400> 50  
Glu Lys Arg Pro Arg Thr Ala Phe Ser Ser Glu Gln Leu Ala Arg Leu  
1 5 10 15

Lys Arg Glu Phe Asn Glu Asn Arg Tyr Leu Thr Glu Arg Arg Arg Gln  
20 25 30

Gln Leu Ser Ser Glu Leu Gly Leu Asn Glu Ala Gln Ile Lys Ile Trp  
35 40 45

Phe Gln Asn Lys Arg Ala Lys Ile Lys Lys Ser Thr  
50 55 60

<210> 51  
<211> 60  
<212> PRT  
<213> Homo sapiens

<400> 51  
Arg Lys Arg Gly Arg Gln Thr Tyr Thr Arg Tyr Gln Thr Leu Glu Leu  
1 5 10 15

Glu Lys Glu Phe His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile  
20 25 30

Glu Ile Ala His Ala Leu Cys Leu Thr Glu Arg Gln Ile Lys Ile Trp  
35 40 45

Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Glu Asn  
50 55 60

<210> 52  
<211> 21  
<212> DNA  
<213> Homo sapiens

<400> 52  
cagaggggac tttccgagag g 21

<210> 53  
<211> 34  
<212> DNA  
<213> HIV

<400> 53  
tacaagggac tttccgctgg ggactttcca gggg 34

<210> 54  
<211> 41  
<212> DNA  
<213> Homo sapiens

<400> 54  
gagaagtgaa agtgggaaat tcctctgaat agagagagga c 41

<210> 55  
<211> 11  
<212> DNA  
<213> Homo sapiens

<400> 55  
ggggactttc c 11

<210> 56  
<211> 11  
<212> DNA  
<213> Homo sapiens

<400> 56  
ggggattccc c 11

<210> 57  
<211> 91  
<212> DNA  
<213> Plasmid pBS-ATG

<400> 57  
agggaaacaaa agcttgcatg cctgcacat ggccatcgat atcgatcccc aattccggcc 60  
cgccccggaa ttgggtaccg agctcgaatt c 91

<210> 58  
<211> 3067  
<212> DNA



<213> Homo sapiens

<220>

<221> CDS

<222> (1175)..(2239)

<223>

<400> 58

```
aacattgcaa ccttataaaa aattaactat ttcgacaatg cgcgagaagg aaattctgtg      60
tttaggtgct ggtgggaaaa cactatctcc agcttgtagg tttgagcatc accagaacca      120
cttgatgaaa tcacacacag aacaagtaga ggaggcaact gtgaatcgtg gggctataaa      180
gccatcaagg gatctgatga aagaaccgac gagacgaacc cccccacccc ccacaacagg      240
atcggcaccc cagagttcaa caagtggctg actttgttaa aacactacgt gggaacccat      300
agtcccgcat cagtagttgc acagccccct ccccgacaga ctacaccgct gtttgctgat      360
ccttgcccac cccatgctct cctcccaggc ccccgttctg ctccctctgtc ctgcggcgct      420
ggattgaacc gcacacaagt ctgcatctgg cacgaattct catgggagcc acgtcatgag      480
gtacgtgggt gcacacctat cacaagaagt cttgcagttc tgactctcct gagctcgggtg      540
ggaaagtctg gatagtacct cccctctcct gccacaaaag cagccctcac attcacaagt      600
ttccaaagca ggtctattga gtttctcttc agagcgagcc tttgtcaaac acacctggag      660
gggggagctc cacctctccc cagcaactca gatcagtgcc ttatttttaa tgctccggcc      720
caatcctgag gtgctgctgg gtttgtgggc tgcgttttgt tgaacctccc cctccccctc      780
ccaacgccct ggcatttgca attaaaactg ggattcaagg gccaaattca agcccagagt      840
gagcagtagg atgtggagct caaagcagag ttgcacctgc tgacccccag cctgaatttg      900
gttcacccag agactacaag tcagaaaggc atgtttagaa agaggcatgc taaggactga      960
tggtggaacg gccaatgtgt ccccaccagc acagtgggga aggctggaca gagaaggaag     1020
agaaggatcc atagagatgt gaaccagaat cagtcgtggt gagctctggg tatatcacta     1080
catgtttaac tcttgcaaga ccgtttgccc agggctttgg taccacaggg ttagagttac     1140
attaaccaca accaccagag aggaactgag gttt atg acc ccc ccc ccc cca aag     1195
```

Met Thr Pro Pro Pro Pro Lys

1

5

```
gtt aga ttt ctg ccg agt ata aag ggg ggg gaa ggg ggg ggt cct tgg      1243
Val Arg Phe Leu Pro Ser Ile Lys Gly Gly Glu Gly Gly Gly Pro Trp
```

10

15

20

```
ttc att tcc ctt cac tgt gtg acc gaa gtt ttg ctt tta ttt gta aac      1291
Phe Ile Ser Leu His Cys Val Thr Glu Val Leu Leu Leu Phe Val Asn
```

25

30

35

```
atc ttg aat tac ccg tgc ttt tcc agt ctt cat cgt gct gtt gtc agg      1339
Ile Leu Asn Tyr Pro Ser Phe Ser Ser Leu His Arg Ala Val Val Arg
```

40

45

50

55

```
cca ctg gag gga att ccc cgt ctc gga acg ccg ccg cca gca cca gca      1387
Pro Leu Glu Gly Ile Pro Arg Leu Gly Thr Pro Pro Pro Ala Pro Ala
```

60

65

70

```
gcc gcg ccg cgc cgc ccc gcc agc tcc gcc gcc atg ctc agc gcc cac      1435
Ala Ala Pro Arg Arg Pro Ala Ser Ser Ala Ala Met Leu Ser Ala His
```

75

80

85

```
cgc ccc gcc gag ccg ccc gcc gtg gag ggc tgc gag ccg ccg cgc aag      1483
Arg Pro Ala Glu Pro Pro Ala Val Glu Gly Cys Glu Pro Pro Arg Lys
```

90

95

100

```
gaa cgg caa ggc ggg ctg ctg ccg ccc gac gac cgc cac gac agc ggg      1531
Glu Arg Gln Gly Gly Leu Leu Pro Pro Asp Asp Arg His Asp Ser Gly
```

105

110

115

```
ctg gac tcc atg aag gag gag gag tac agg cag ctg gtg cgg gag ctg      1579
Leu Asp Ser Met Lys Glu Glu Glu Tyr Arg Gln Leu Val Arg Glu Leu
```

120

125

130

135

```
gag gac atc cgc ctg cag ccc cgc gag ccg ccc gcc cgg ccg cac gcc      1627
Glu Asp Ile Arg Leu Gln Pro Arg Glu Pro Pro Ala Arg Pro His Ala
```

140

145

150

tgg gcc cag cag ctc acc gag gac ggc gac act ttt ctc cac ttg gcg	1675
Trp Ala Gln Gln Leu Thr Glu Asp Gly Asp Thr Phe Leu His Leu Ala	
155 160 165	
atc att cac gag gaa aag gcc ctg agc ctg gag gtg atc cgg cag gcc	1723
Ile Ile His Glu Glu Lys Ala Leu Ser Leu Glu Val Ile Arg Gln Ala	
170 175 180	
gct ggg gac gcc gcc ttc ctg aac ttc cag aac aac ctc agc cag act	1771
Ala Gly Asp Ala Ala Phe Leu Asn Phe Gln Asn Asn Leu Ser Gln Thr	
185 190 195	
ccg ctc cac ctg gcg gtg atc acg gac cag gcc gaa atc gcc gag cac	1819
Pro Leu His Leu Ala Val Ile Thr Asp Gln Ala Glu Ile Ala Glu His	
200 205 210 215	
ctg ctg aag gct ggc tgc gac ctg gat gtc agg gac ttc cgt ggg aac	1867
Leu Leu Lys Ala Gly Cys Asp Leu Asp Val Arg Asp Phe Arg Gly Asn	
220 225 230	
acc ccg ctc cac atc gcc tgc cag cag ggc tgc ctc cgc agc gtc agt	1915
Thr Pro Leu His Ile Ala Cys Gln Gln Gly Ser Leu Arg Ser Val Ser	
235 240 245	
gtc ctc acg cag cac tgc cag ccc cac cac ctc ctc gcc gtc ctg cag	1963
Val Leu Thr Gln His Cys Gln Pro His His Leu Leu Ala Val Leu Gln	
250 255 260	
gcc acc aac tac aac ggc cat aca tgt ctc cat ttg gca tct att caa	2011
Ala Thr Asn Tyr Asn Gly His Thr Cys Leu His Leu Ala Ser Ile Gln	
265 270 275	
gga tac ctg gct gtt gtc gaa tac ctg ctg tcc tta gga gca gat gta	2059
Gly Tyr Leu Ala Val Val Glu Tyr Leu Leu Ser Leu Gly Ala Asp Val	
280 285 290 295	
aat gct cag gag cca tgc aat ggg aga aca gca cta cac ttg gcc gta	2107
Asn Ala Gln Glu Pro Cys Asn Gly Arg Thr Ala Leu His Leu Ala Val	
300 305 310	
gac ctt cag aac tca gac ctg gtg tca ctt ctg gtg aaa cac ggg cca	2155
Asp Leu Gln Asn Ser Asp Leu Val Ser Leu Leu Val Lys His Gly Pro	
315 320 325	
gat gtg aac aaa gtg acc tac cag ggc tac tcc cca tac cag ctt aca	2203
Asp Val Asn Lys Val Thr Tyr Gln Gly Tyr Ser Pro Tyr Gln Leu Thr	
330 335 340	
tgg gca gag aca acg cca gca tac agg agc agc tga agctgctgac	2249
Trp Ala Glu Thr Thr Pro Ala Tyr Arg Ser Ser	
345 350	
cacagctgac ctgcagatac tgcccgaag tgaggatgag gagagcagtg aatcagagcc	2309
agagttcaca gaggatgaac ttatgtatga tgactgctgt attggaggaa gacagctgac	2369
attttaaagc agaggtttct gtgagaagtg actgtgtaca tatgtatagg aaaaaaagcc	2429
tgacttttctt catttaaaaa gaaagtctat actcgaagga gaaaaaagta ctgagatact	2489
acactgcccc gccaggagca catcatccta acagggtcca tgctctgacc tgtacttaag	2549
taacgggatg ggatgtgtaa catcgtaaag agatcagtga acatgcacac catctgataa	2609
agagccacgt tatctaattt ctctgccaca tgaggataac ggactgcacg tccaatgtgc	2669
tgttgtcaga aatgcgtttg agagctgcct tgtgacacta agtgctgtga ggagtgtca	2729
tccccctcgg tggcaagaca ggcttgacaa aacgtcccat ctgcttgaag actgtgaggt	2789
tggtcattagg ttgaggcact gctgtgccct gctccctgac cctggctgct cagggttgag	2849
gagtcggacc atgggagagg tgacctggct gctgggagga aggtagcaat gatgttaact	2909
gtgggcattt ggaaactgtg tgtttcacac catgtgtgtc ataattgcta cactttttag	2969
caactgtata gaatgtaaat actgtacatc tttgtttata attattttgg tacctgtgag	3029
atatgtattt attaaaaaag gcagatttct gtaaaaaa	3067

<210> 59  
 <211> 354  
 <212> PRT

<213> Homo sapiens

<400> 59

Met	Thr	Pro	Pro	Pro	Pro	Lys	Val	Arg	Phe	Leu	Pro	Ser	Ile	Lys	Gly	
1				5				10						15		
Gly	Glu	Gly	Gly	Gly	Pro	Trp	Phe	Ile	Ser	Leu	His	Cys	Val	Thr	Glu	
			20					25					30			
Val	Leu	Leu	Leu	Phe	Val	Asn	Ile	Leu	Asn	Tyr	Pro	Ser	Phe	Ser	Ser	
		35					40					45				
Leu	His	Arg	Ala	Val	Val	Arg	Pro	Leu	Glu	Gly	Ile	Pro	Arg	Leu	Gly	
	50					55					60					
Thr	Pro	Pro	Pro	Ala	Pro	Ala	Ala	Ala	Pro	Arg	Arg	Pro	Ala	Ser	Ser	
65					70					75					80	
Ala	Ala	Met	Leu	Ser	Ala	His	Arg	Pro	Ala	Glu	Pro	Pro	Ala	Val	Glu	
				85					90					95		
Gly	Cys	Glu	Pro	Pro	Arg	Lys	Glu	Arg	Gln	Gly	Gly	Leu	Leu	Pro	Pro	
			100					105					110			
Asp	Asp	Arg	His	Asp	Ser	Gly	Leu	Asp	Ser	Met	Lys	Glu	Glu	Glu	Tyr	
		115					120					125				
Arg	Gln	Leu	Val	Arg	Glu	Leu	Glu	Asp	Ile	Arg	Leu	Gln	Pro	Arg	Glu	
	130					135					140					
Pro	Pro	Ala	Arg	Pro	His	Ala	Trp	Ala	Gln	Gln	Leu	Thr	Glu	Asp	Gly	
145					150					155					160	
Asp	Thr	Phe	Leu	His	Leu	Ala	Ile	Ile	His	Glu	Glu	Lys	Ala	Leu	Ser	
				165					170					175		
Leu	Glu	Val	Ile	Arg	Gln	Ala	Ala	Gly	Asp	Ala	Ala	Phe	Leu	Asn	Phe	
		180						185					190			
Gln	Asn	Asn	Leu	Ser	Gln	Thr	Pro	Leu	His	Leu	Ala	Val	Ile	Thr	Asp	
		195					200					205				
Gln	Ala	Glu	Ile	Ala	Glu	His	Leu	Leu	Lys	Ala	Gly	Cys	Asp	Leu	Asp	
	210					215					220					
Val	Arg	Asp	Phe	Arg	Gly	Asn	Thr	Pro	Leu	His	Ile	Ala	Cys	Gln	Gln	
225					230					235					240	
Gly	Ser	Leu	Arg	Ser	Val	Ser	Val	Leu	Thr	Gln	His	Cys	Gln	Pro	His	
				245					250					255		
His	Leu	Leu	Ala	Val	Leu	Gln	Ala	Thr	Asn	Tyr	Asn	Gly	His	Thr	Cys	
			260					265					270			
Leu	His	Leu	Ala	Ser	Ile	Gln	Gly	Tyr	Leu	Ala	Val	Val	Glu	Tyr	Leu	
		275					280					285				

Leu Ser Leu Gly Ala Asp Val Asn Ala Gln Glu Pro Cys Asn Gly Arg  
290 295 300

Thr Ala Leu His Leu Ala Val Asp Leu Gln Asn Ser Asp Leu Val Ser  
305 310 315 320

Leu Leu Val Lys His Gly Pro Asp Val Asn Lys Val Thr Tyr Gln Gly  
325 330 335

Tyr Ser Pro Tyr Gln Leu Thr Trp Ala Glu Thr Thr Pro Ala Tyr Arg  
340 345 350

Ser Ser

<210> 60  
<211> 11  
<212> DNA  
<213> Homo sapiens

<400> 60  
tgggattcc ca